

This manual includes

GOAL

Customizable, versatile
online exam question bank.
Thousands of questions!

Access your exclusive
StudyPlus⁺ bonus content:

GOAL | Career Guides | Samples

* Key Code Inside *



Exam MAS-I Study Manual



Fall 2019 Edition, Volume I

Ambrose Lo, FSA, Ph.D., CERA

ACTEX Learning | Learn Today. Lead Tomorrow.



Exam MAS-I Study Manual

Fall 2019 Edition, Volume I

Ambrose Lo, FSA, Ph.D., CERA



Actuarial & Financial Risk Resource Materials
Since 1972

Copyright © 2019 SRBooks, Inc.

ISBN: 978-1-63588-852-2

Printed in the United States of America.

No portion of this ACTEX Study Manual may be reproduced or transmitted in any part or by any means without the permission of the publisher.

Contents

Preface	vii
I Probability Models (Stochastic Processes and Survival Models)	1
1 Poisson Processes	3
1.1 Fundamental Properties	5
Problems	18
1.2 Hazard Rate Function	27
Problems	48
1.3 Inter-arrival and Waiting Distributions	73
Problems	83
1.4 New Poisson Processes from Old Ones	90
1.4.1 Thinning	90
1.4.2 Sums of Poisson Processes	104
1.4.3 Problems	106
1.5 Compound Poisson Processes	122
Problems	134
2 Reliability Theory	149
2.1 Typical Systems	150
2.2 Reliability of Systems of Independent Components	157
2.3 Expected System Lifetime	172
2.4 End-of-chapter Problems	181
3 Markov Chains	199
3.1 Transition Probabilities	200
3.1.1 Fundamentals	200
3.1.2 Chapman–Kolmogorov Equations	207
3.1.3 [HARDER!] Path-dependent Probabilities	212
3.1.4 Problems	218
3.2 The Gambler’s Ruin Problem	253
Problems	261
3.3 Classification of States	267
Problems	275
3.4 Long-Run Proportions and Limiting Probabilities	277
Problems	291
3.5 Mean Time Spent in Transient States	318

Problems	324
3.6 Branching Processes	333
Problems	338
3.7 Time Reversible Markov Chains	340
4 Elementary Life Contingencies	343
4.1 Life Tables	344
4.2 Life Insurances	350
4.3 Life Annuities	357
4.4 Net Level Premiums	366
4.5 End-of-chapter Problems	372
5 Simulation	397
5.1 Inversion Method	397
5.2 Acceptance-Rejection Method	403
5.3 End-of-chapter Problems	408
II Statistics	419
6 Claim Severity Distributions	423
6.1 Continuous Distributions for Modeling Claim Severity	423
6.2 Mixture Distributions	428
6.3 Tail Properties of Claim Severity	430
6.4 Effects of Coverage Modifications	433
6.4.1 Deductibles	433
6.4.2 Policy Limits	442
6.4.3 Coinsurance and Combinations of Coverage Modifications	445
6.5 End-of-chapter Problems	453
7 Parametric Estimation	477
7.1 Quality of Estimators	478
Problems	490
7.2 Some Parametric Estimation Methods	509
7.2.1 Method of Moments	509
7.2.2 Method of Percentile Matching	515
7.2.3 Problems	520
7.3 Maximum Likelihood Estimation	534
7.3.1 Determining MLEs by Calculus	535
7.3.2 Distributions Which Require Special Techniques	544
7.3.3 Remarks on MLE	550
7.3.4 Problems	559
7.4 Variance of Maximum Likelihood Estimators	600
Problems	610
7.5 Sufficiency	622
Problems	630

8 Hypothesis Testing	641
8.1 Fundamental Ideas of Hypothesis Testing	642
Problems	654
8.2 Test for Means of Normal Populations	673
8.2.1 One-sample Case	673
8.2.2 Two-sample Case	680
8.2.3 Problems	689
8.3 Test for Variances of Normal Populations	702
8.3.1 One-sample Case	702
8.3.2 Two-sample Case	707
8.3.3 Problems	717
8.4 Tests for Binomial Populations	725
8.4.1 One-sample Case	725
8.4.2 Two-sample Case	731
8.4.3 Problems	737
8.5 Most Powerful and Uniformly Most Powerful Tests	746
Problems	758
8.6 Likelihood Ratio Tests	768
Problems	782
8.7 Chi-Square Tests for Count Data	787
8.7.1 Chi-Square Goodness-of-Fit Test	787
8.7.2 Chi-Square Test of Independence	796
8.7.3 Problems	804
8.8 Kolmogorov–Smirnov Test	823
Problems	827
9 Nonparametric Estimation	835
9.1 Kernel Density Estimation	835
9.1.1 Basic Ideas	835
9.1.2 Case Study 1: Rectangular Kernel	836
9.1.3 Case Study 2: Triangular and Gaussian Kernels	840
9.1.4 Problems	844
9.2 Order Statistics	849
9.2.1 Basics of Order Statistics	849
9.2.2 Two Important Case Studies	863
9.2.3 Joint Distribution of Order Statistics	870
9.2.4 Problems	872
III Extended Linear Models	897
10 Linear Regression Models	899
10.1 Model Formulation and Parameter Estimation	900
Problems	910
10.2 Model Construction	922
10.2.1 Types of Predictors	922
10.2.2 Interaction	929
10.2.3 Problems	934

10.3	Statistical Inference	940
10.3.1	Properties of Least Squares Estimators	940
10.3.2	t -test for the Significance of a Single Regression Coefficient	950
10.3.3	F -test for the Significance of Several Regression Coefficients	955
10.3.4	Problems	965
10.4	ANOVA	999
10.4.1	One-factor ANOVA	999
10.4.2	Two-factor ANOVA	1006
10.4.3	Problems	1010
10.5	Model Diagnostics	1016
10.5.1	High Leverage Points	1016
10.5.2	Outliers	1018
10.5.3	Overall Influence Measures	1019
10.5.4	Other Diagnostic Tools	1025
	Problems	1032
11	Statistical Learning	1037
11.1	A Primer on Statistical Learning	1038
11.1.1	Fundamental Concepts	1038
11.1.2	Assessing the Accuracy of a Model	1042
11.2	Resampling Methods	1047
11.2.1	Validation Set Approach	1047
11.2.2	Cross-Validation	1048
11.2.3	Bootstrap	1056
11.3	Variable Selection	1058
11.3.1	Model Comparison Statistics	1058
11.3.2	Best Subset Selection	1063
11.3.3	Automatic Variable Selection Procedures	1066
11.3.4	Using Validation and Cross-Validation for Variable Selection	1075
11.4	Shrinkage Methods	1077
11.4.1	Ridge Regression	1077
11.4.2	Lasso	1081
11.5	Dimension Reduction Methods	1090
11.5.1	Principal Components Analysis	1090
11.5.2	Applications of PCA to Supervised Learning	1095
11.6	Nonlinear Models	1099
11.6.1	Regression Splines	1099
11.6.2	Smoothing Splines	1104
11.6.3	Local Regression	1108
11.6.4	Generalized Additive Models	1109
11.7	End-of-chapter Problems	1111
12	Generalized Linear Models	1139
12.1	Exponential Family of Distributions	1140
12.2	Estimation of Parameters	1147
12.3	Prediction	1154
12.4	Assessing Model Fit by Deviance	1160

12.5 Hypothesis Tests on Regression Coefficients	1166
12.6 End-of-chapter Problems	1174
13 Special GLMs for Insurance Data	1193
13.1 Logistic Regression	1194
13.2 Nominal and Ordinal Logistic Regression	1216
13.3 Poisson Regression	1223
13.4 End-of-chapter Problems	1233
IV Time Series with Constant Variance	1257
14 Introduction to Time Series Analysis	1261
14.1 Trends and Seasonality	1261
14.2 Autocorrelation and Cross Correlation	1267
14.3 End-of-chapter Problems	1277
15 ARIMA Models	1289
15.1 White Noise and Random Walk Processes	1289
15.2 Autoregressive Models	1294
15.3 Moving Average Models	1312
15.4 ARMA Models	1319
15.5 Non-stationary Models	1324
15.6 End-of-chapter Problems	1330
V Practice Examinations	1355
Practice Exam 1	1358
Solutions to Practice Exam 1	1382
Practice Exam 2	1398
Solutions to Practice Exam 2	1423
Practice Exam 3	1437
Solutions to Practice Exam 3	1462

Preface (Please Read Carefully!)

Exam MAS-I (Modern Actuarial Statistics I) is a relatively new exam which was introduced in Spring 2018 by the Casualty Actuarial Society (CAS). It is offered two times a year. In Fall 2019, the exam will take place on **October 24, 2019 (Thursday)**. The registration deadline is August 30, 2019. This exam replaces its predecessor Exam S (*Statistics and Probabilistic Models*), which is a short-lived exam offered only five times, from Fall 2015 to Fall 2017. Exam S, in turn, was developed from the old Exam LC (*Models for Life Contingencies*), Exam ST (*Models for Stochastic Processes and Statistics*), and Exam 3L (*Life Contingencies and Statistics*). The introduction of Exam MAS-I is in response to the discontinuation of Exam C/4 in July 2018, which the CAS saw as an opportunity to revamp Exam S with a heavier focus on contemporary statistical methods as a way to enhance the statistical literacy of property and casualty actuaries in this day and age. You will considerably sharpen your statistics toolkit as a result of taking (and, with the use of this study manual, passing!) Exam MAS-I.

Syllabus

The syllabus of Exam MAS-I, available from <http://www.casact.org/admissions/syllabus/ExamMASI.pdf>, is *extremely broad* in scope (that explains the sheer size of this manual!), covering miscellaneous topics in applied probability, mathematical statistics, applied statistical modeling and time series analysis, many of which are new topics not tested in any SOA/CAS past exams. As a rough estimate, you need at least *three months* of intensive study to master the material in this exam.¹ The specific sections of the syllabus along with their approximate weights in the exam are shown below:

Section	Range of Weight
A. Probability Models (Stochastic Processes & Survival Models)	20–35%
B. Statistics	15–30%
C. Extended Linear Models	30–50%
D. Time Series with Constant Variance	10–20%

Compared with the former Exam S, both Sections C and D have enjoyed a heavier weight, from 25–40% to 30–50% and from 5–10% to 10–20%, respectively. Sections A, B, and D are more or less taken intact from the syllabuses of Exam S, ST, LC, and 3L. As a result, you can find lots of relevant past exam questions on these three sections. Section C has experienced the biggest change, with a new textbook on statistical learning added.

¹It is true that one need not master every topic in order to pass this exam.

Exam Format

Exam MAS-I is a four-hour multiple-choice exam. According to the list of MAS-I frequently asked questions (http://www.casact.org/cms/files/New_CAS_Exams_MAS_I_and_II_FAQs_1.pdf), the exam will consist of approximately 35 to 40 questions. However, the Spring 2019, Fall 2018, and Spring 2018 MAS-I exams all have 45 questions and you can expect future exams to have 45 questions as well. Each correct answer receives 2 points with the total score being $45 \times 2 = 90$ while *0.5 points will be subtracted for each incorrect answer*, unlike other multiple-choice exams you have taken. Before the start of the exam, there will be a fifteen-minute reading period in which you can silently read the questions and check the exam booklet for missing or defective pages. However, writing will not be permitted during this time, neither will the use of calculators.

Given the similarity between Exam MAS-I and Exam S (the syllabus of the latter is available from <http://www.casact.org/admissions/syllabus/ExamS.pdf>) in terms of their structure and topics, we may use Exam S as a rough proxy for Exam MAS-I. The following table shows the distribution of exam questions from Fall 2015 to Spring 2019, categorized into the four sections:

Section	Number of Questions							
	15F	16S	16F	17S	17F	18S	18F	19S
A. Probability Models	13	12	16	15	16	14	14	14
B. Statistics	17	15	14	13	14	9	11	12
C. Extended Linear Models	11	15	11	14	12	17	15	14
D. Time Series with Constant Variance	4	3	4	3	3	5	5	5
Total	45	45	45	45	45	45	45	45

You can see that roughly the same number of exam questions were set on Sections A and C, although Section C is proclaimed to be the most important section (perhaps even the examiners found it hard to set questions on this section!). According to <https://www.casact.org/admissions/passmarks/examMAS-I.pdf>, the pass marks for the Fall 2018 and Spring 2018 MAS-I examsⁱⁱ were both **53 out of 90**, which means that candidates needed to answer at least **27 out of 45 questions** correctly to earn a pass (ignoring the guessing penalty).

Here are the characteristics of a typical CAS multiple-choice exam:

1. The questions are almost always arranged in the same order as the topics in the exam syllabus, so Question 1 is almost always a Poisson process question and Question 45 is almost always a time series question. This implicitly gives you a hint as to which topic an exam question is testing. (Within the same section of the syllabus, however, the exam questions can appear in a rather erratic order.)
2. A number of exam questions bear a striking resemblance to recent CAS exam questions, sometimes even with the same numerical values. In the Spring 2019 Exam MAS-I, for example,

Question 21 was adapted from Question 11 of the Spring 2015 Exam ST,
 Question 26 was adapted from Question 18 of the Fall 2015 Exam ST,
 Questions 31 and 32 were based on Question 35 of the Fall 2018 Exam MAS-I,
 Question 39 was taken directly from Question 39 of the Fall 2018 Exam MAS-I.

This attests to the importance of practicing numerous past exam problems, an abundance of which are carefully discussed and solved in this study manual.

ⁱⁱThe pass mark for the Spring 2019 Exam MAS-I is not yet available when this manual goes in press.

3. The scope of an exam can be narrow at times with several questions testing essentially the same topic in much the same way. In the Spring 2019 Exam MAS-I, for example, Questions 3 and 6 both test the variance of the exponential loss in the presence of a deductible, Questions 18 and 19 both test the estimation surrounding the uniform distribution, and Questions 21 and 22 both test the concept of Type I error.
4. Most answer choices are in the form of ranges, e.g.:
 - A. Less than 1%
 - B. At least 1%, but less than 2%
 - C. At least 2%, but less than 3%
 - D. At least 3%, but less than 4%
 - E. At least 4%

If your answer is much lower than the bound indicated by Answer A or much higher than that suggested by Answer E, do check your calculations. Chances are that you have made computational mistakes, but this is not definitely the case (sometimes the CAS examiners themselves made a mistake!).

As mentioned above, 0.5 points will be deducted for each incorrect answer, so unless you can eliminate two or three of the answer choices, it will be wise of you not to answer questions which you have completely no idea by pure guesswork.

What is Special about This Study Manual?

We fully understand that you have an acutely limited amount of time for study and that the exam syllabus is insanely broad. With this in mind, the overriding objective of this study manual is to help you grasp the material in Exam MAS-I, which is a new and challenging exam, effectively and efficiently, and *pass it with considerable ease*. Here are some of the invaluable features of this manual for achieving this all-important goal:

- Each chapter and section starts by explicitly stating which learning objectives and outcomes of the MAS-I exam syllabus we are going to cover, to assure you that we are on track and hitting the right target.
- The knowledge statements of the syllabus are then demystified by precise and concise expositions synthesized from the syllabus readings, helping you acquire a deep and solid understanding of the subject matter.
- Formulas and results of utmost importance are boxed for easy identification and memorization.
- Mnemonics and shortcuts are emphasized, so are highlights of important exam items and common mistakes committed by students.
- While the focus of this study manual is on exam preparation, we will not shy away from explaining the meaning of various formulas in the syllabus. The interpretations and insights provided will foster a genuine understanding of the syllabus material and discourage slavish

memorization. At times, we will present brief derivations with the aim of helping you appreciate the structure of the formulas in question. It is the author's belief and personal experience that a solid understanding of the underlying concepts is always conducive to achieving good exam results.

- To succeed in any actuarial exam, the importance of practicing a wide variety of non-trivial problems to sharpen your understanding and to develop proficiency, as always, cannot be overemphasized. This study manual embraces this learning by doing approach and intersperses its expositions with more than **540 in-text examples** and **820 end-of-chapter/section problems** (the harder ones are labeled as [**HARDER!**] or [**VERY HARD!!**]), which are original or taken from relevant SOA/CAS past exams, all with step-by-step solutions and problem-solving remarks, to consolidate your understanding and to give you a sense of what you can expect to see in the real exam. As you read this manual, skills are honed and confidence is built. As a general guide, you should *study all of the in-text examples paying particular attention to recent Exam MAS-I/S questions* and *work out at least all of the past CAS problems included in the end-of-chapter/section problem sets*. Of course, the more problems you do, the better.
- Three full-length practice exams updated for the Spring 2019 MAS-I exam syllabus and designed to mimic the real exam conclude this study manual giving you a holistic review of the syllabus material.

New to the Fall 2019 Edition

- While there have been updates throughout the entire manual in terms of content, clarity, and exam focus in response to the recent MAS-I exams, the improvements are particularly significant in Sections 6.4, 8.3, 10.1, 10.2, 10.3, 11.2, 11.3, and 11.6. A number of new examples and end-of-chapter/section problems have been added. *In Bayesian parlance, we learn from experience as life moves on!*
- Compared to the Spring 2019 edition, this new edition of the manual has more than 40 new in-text examples and 40 end-of-chapter/section problems. A number of these new examples and problems are of the true-or-false type, whose popularity has risen in recent MAS-I exams.
- All of the 45 questions from the very recent Spring 2019 Exam MAS-I are inserted into appropriate sections of this manual and carefully discussed and solved. Variants of some of these exam questions are developed.
- All known typographical errors have been fixed.

Exam Tables

In the exam, you will be supplied with a variety of tables, including:

- *Standard normal distribution table* (used throughout this study manual)

You will need this table for values of the standard normal distribution function or standard normal quantiles, when you work with normally distributed random variables or perform normal approximation.

- *Illustrative Life Table* (used mostly in Chapter 4 of this study manual)

You will need this when you are told that mortality of the underlying population follows the Illustrative Life Table.

- *A table of distributions for a number of common continuous and discrete distributions and the formulas for their moments and other probabilistic quantities* (used throughout Parts I and II of this study manual)

This big table provides a great deal of information about some common (e.g., exponential, gamma, lognormal, Pareto) as well as less common distributions (e.g., inverse exponential, inverse Gaussian, Pareto, Burr, etc.). When an exam question centers on these distributions and quantities such as their means or variances are needed, consult this table.

- *Quantiles of t -distribution, F -distribution, chi-square distribution* (used in Chapters 8, 10, 12 and 13 of this study manual)

These quantiles will be of use when you perform parametric hypothesis tests.

You should download these tables from http://www.casact.org/admissions/syllabus/MASI_Tables.pdf right away, print out a copy and learn how to locate the relevant entries in these tables because they will be intensively used during your study as well as in the exam.

Acknowledgment

I would like to thank my colleagues, Professor Elias S. W. Shiu and Dr. Michelle A. Larson, at the University of Iowa for sharing with me many pre-2000 SOA/CAS exam papers. These hard-earned old exam papers have proved invaluable in illustrating a number of less commonly tested exam topics. Thanks are also due to Mr. Zhaofeng Tang, doctoral student in Actuarial Science at the University of Iowa, for his professional assistance in the production of some of the graphs in this study manual.

Errata

While we go to great lengths to polish and proofread this manual, some mistakes will inevitably go unnoticed. The author wishes to apologize in advance for any errors, typographical or otherwise, and would greatly appreciate it if you could bring them to his attention by *sending any errors you identify to ambrose-lo@uiowa.edu and c.c. support@actexamdriver.com*. Compliments and criticisms are also welcome. The author will try his best to respond to any inquiries as soon as possible and an ongoing errata list will be maintained online at <https://sites.google.com/site/ambroseloyp/publications/MAS-I>. Students who report errors will be entered into a quarterly drawing for a \$100 in-store credit.

Ambrose Lo
May 2019
Iowa City, IA

About the Author

Ambrose Lo, FSA, CERA, is currently Associate Professor of Actuarial Science with tenure at the Department of Statistics and Actuarial Science at the University of Iowa. He received his Ph.D. in Actuarial Science from the University of Hong Kong in 2014, with dependence structures, risk measures, and optimal reinsurance being his research interests, and joined the University of Iowa in August 2014. His research papers have been published in top-tier actuarial journals, such as *ASTIN Bulletin: The Journal of the International Actuarial Association*, *Insurance: Mathematics and Economics*, and *Scandinavian Actuarial Journal*. He has taught courses on financial derivatives, mathematical finance, life contingencies, credibility theory, advanced probability theory, and regression and time series analysis. His emphasis in teaching is always placed on the development of a thorough understanding of the subject matter complemented by concrete problem-solving skills. Besides coauthoring the *ACTEX Study Manual for SOA Exam SRM* (Fall 2019 Edition), he is also the sole author of the *ACTEX Study Manual for CAS Exam MAS-I* (Fall 2019 Edition), *ACTEX Study Manual for SOA Exam PA* (December 2019 Edition) and the textbook *Derivative Pricing: A Problem-Based Primer* (2018) published by Chapman & Hall/CRC Press.

Part I

Probability Models (Stochastic Processes & Survival Models)

Chapter 1

Poisson Processes

LEARNING OBJECTIVES

1. Understand and apply the properties of Poisson processes:

- For increments in the homogeneous case
- For interval times in the homogeneous case
- For increments in the non-homogeneous case
- Resulting from special types of events in the Poisson process
- Resulting from sums of independent Poisson processes

Range of weight: 0-5 percent

2. For any Poisson process and the inter-arrival and waiting distributions associated with the Poisson process, calculate:

- Expected values
- Variances
- Probabilities

Range of weight: 0-5 percent

3. For a compound Poisson process, calculate moments associated with the value of the process at a given time.

Range of weight: 0-5 percent

4. Apply the Poisson Process concepts to calculate the hazard function and related survival model concepts.

- Relationship between hazard rate, probability density function and cumulative distribution function
- Effect of memoryless nature of Poisson distribution on survival time estimation

Range of weight: 2-8 percent

Chapter overview: As a prospective P&C actuary, you would be interested in monitoring the number of insurance claims an insurance company receives as time goes by and how these claims can be appropriately analyzed by means of sound statistical analysis. In Exam MAS-I, we shall learn one way of modeling the flow of insurance claims – the Poisson process.

This part of the syllabus has two required readings:

- (1) *A study note by J.W. Daniel*

The study note is precise and concise, introducing main results mostly without proof and supplementing its exposition with a few simple examples. It is suitable for a first-time introduction to Poisson processes.

- (2) *The book entitled Introduction to Probability Models by S.M. Ross.*

This is a textbook used by a number of college courses on elementary applied probability. It balances rigor and intuition, and presents the theory of Poisson processes at a level that is much deeper than that in the study note by Daniel. In particular, it treats the conditional distribution of the arrival times as well as the interplay between two independent Poisson processes. A conspicuous feature of this book is its large number of sophisticated examples and exercises which require a large amount of ingenuity and cannot be done in a reasonable exam setting. This study manual improves the practicality of the book and rephrases these otherwise intractable examples in an exam tone.

You can download this book (Eleventh Edition) “legally” from *ScienceDirect* via the following link, chapter by chapter, if your university has subscribed to it:

<http://www.sciencedirect.com/science/book/9780124079489>

You should do so because the book has a number of good exercises which will be solved in full in this study manual (the questions cannot be reproduced here because of copyright issues). The exercises had been the theme of some past SOA/CAS examination questions, so you should not despise these exercises as irrelevant and useless.

The Daniel study note has been on the syllabuses of Exams 3, 3L and ST, whereas *Introduction to Probability Models* just entered the syllabus of Exam S in Fall 2015. As a result of the addition of the latter reading, we expect more complex exam questions on Poisson processes in Exam MAS-I. In total, expect about 4 questions on the material of the entire chapter.

1.1 Fundamental Properties

MAS-I KNOWLEDGE STATEMENT(S)

- 1a. Poisson process
- 1b. Non-homogeneous Poisson process

OPTIONAL SYLLABUS READING(S)

- Ross, Subsections 5.3.1 and 5.3.2
- Daniel, Section 1.1 and Subsection 1.4.1

Definition. By definition, a *Poisson process* $\{N(t), t \geq 0\}$ with *rate function* (also known as *intensity function*) $\lambda(\cdot)$ ⁱ is a stochastic process, namely, a collection of random variables indexed by time t (in an appropriate unit, e.g., minute, hour, month, year, etc.), satisfying the following properties:

1. (*Counting*) $N(0) = 0$, $N(t)$ is non-decreasing in t and takes non-negative integer values only.

Interpretation: $N(t)$ counts the number of claims which are submitted on or before time t . Thus $N(0)$ is 0 (we assume that should be no claims before the insurance company starts its business), $N(t)$ cannot decrease in time and must be integer-valued.

2. (*Distribution of increments are Poisson random variables*) For $s < t$, the increment $N(t) - N(s)$, which counts the number of events in the interval $(s, t]$, is a Poisson random variable with mean $\Lambda = \int_s^t \lambda(y) dy$.

Interpretation: Increments of a Poisson process, as its name suggests, are Poisson random variables with mean computed by integrating the rate function over the same interval. In this regard, we can see that the rate function of a Poisson process completely specifies the distribution of each increment.

3. (*Increments are independent*) If $(s_1, t_1]$ and $(s_2, t_2]$ are non-overlapping intervals, then $N(t_1) - N(s_1)$ and $N(t_2) - N(s_2)$ are independent random variables.

Interpretation: This is the most amazing property of a Poisson process. Its increments not only follow Poisson distribution, but also are independent on disjoint intervals (e.g., $(0, 1)$ and $(2, 5)$ are disjoint intervals, so are $(3, 4]$ and $(4, 5]$). This means that, in this model, the frequency of claims you received last month has nothing to do with the frequency this month.

In the context of insurance applications, we interpret $N(t)$ as the number of claims that occur on or before time t . The same interpretation can easily carry over to more general contexts where we are interested in counting a particular type of event, e.g., the number of customers that enter a store, the number of cars passing through an intersection, the number of lucky candidates passing Exam MAS-I, etc.

ⁱThe study note by Daniel simply writes a Poisson process as N in short. While this is a perfectly correct way of writing, some students may confuse that with a Poisson random variable N . Also, here we write $\lambda(\cdot)$ with a parenthesis containing the argument of the function instead of just λ to emphasize that $\lambda(\cdot)$ is a function.

Homogeneous Poisson processes. A Poisson process whose rate function is constant, say $\lambda(t) = \lambda$ for all $t \geq 0$, is called a *homogeneous* Poisson process. In addition to having independent increments, a homogeneous Poisson process also possesses *stationary increments*, meaning that the distribution of $N(t + s) - N(s)$ depends only on the length of the interval, which is t in this case, but not on s .

Probability calculations. The second and third properties of a Poisson process allow us to calculate many probabilistic quantities, such as the probability of a certain number of events, as well as the expected and variance of the number of events in a particular time interval. These two properties will be intensively used in exam questions. The following string of past exam questions serves as excellent illustrations.

RECALL FROM EXAM P/1

Just in case you forgot:

1. The probability mass function of a Poisson random variable X with parameter λ (a scalar, not a function) is given by

$$\Pr(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

The mean and variance of X are both equal to λ .

2. If X_1, X_2, \dots, X_n are independent Poisson random variables with respective means $\lambda_1, \lambda_2, \dots, \lambda_n$, then $X_1 + X_2 + \dots + X_n$ is also a Poisson random variable with a mean of $\lambda_1 + \lambda_2 + \dots + \lambda_n$. In other words, the sum of independent Poisson random variables is also a Poisson random variable whose mean is the sum of the individual Poisson means.

Example 1.1.1. (SOA Exam P Sample Question 173: Warm-up question) In a given region, the number of tornadoes in a one-week period is modeled by a Poisson distribution with mean 2. The numbers of tornadoes in different weeks are mutually independent.

Calculate the probability that fewer than four tornadoes occur in a three-week period.

- A. 0.13
- B. 0.15
- C. 0.29
- D. 0.43
- E. 0.86

Ambrose's comments: This is not even a Poisson process question. It simply reminds you of how probabilities for a Poisson random variable are typically calculated.

Solution. We are interested in $\Pr(N_1 + N_2 + N_3 < 4)$, where N_i is the number of tornadoes in the i^{th} week for $i = 1, 2, 3$. As $N_1 + N_2 + N_3$ is also a Poisson random variable with a mean of $3(2) = 6$, we have

$$\begin{aligned}\Pr(N_1 + N_2 + N_3 < 4) &= \sum_{i=0}^3 \Pr(N_1 + N_2 + N_3 = i) \\ &= e^{-6} \left(1 + 6 + \frac{6^2}{2!} + \frac{6^3}{3!} \right) \\ &= \boxed{0.1512}. \quad (\text{Answer: B})\end{aligned}$$

□

Example 1.1.2. (CAS Exam MAS-I Spring 2018 Question 3: Probability – I) The number of cars passing through the Lexington Tunnel follows a Poisson process with rate:

$$\lambda(t) = \begin{cases} 16 + 2.5t & \text{for } 0 < t \leq 8 \\ 52 - 2t & \text{for } 8 < t \leq 12 \\ -20 + 4t & \text{for } 12 < t \leq 18 \\ 160 - 6t & \text{for } 18 < t \leq 24 \end{cases}$$

Calculate the probability that exactly 50 cars pass through the tunnel between times $t = 11$ and $t = 13$.

- A. Less than 0.01
- B. At least 0.01, but less than 0.02
- C. At least 0.02, but less than 0.03
- D. At least 0.03, but less than 0.04
- E. At least 0.04

Solution. The number of cars between $t = 11$ and $t = 13$ is a Poisson random variable with parameter

$$\begin{aligned}\int_{11}^{13} \lambda(t) dt &= \int_{11}^{12} (52 - 2t) dt + \int_{12}^{13} (-20 + 4t) dt \\ &= [52t - t^2]_{11}^{12} + [-20t + 2t^2]_{12}^{13} \\ &= 59.\end{aligned}$$

Hence the probability of exactly 50 cars between $t = 11$ and $t = 13$ is

$$\frac{e^{-59} 59^{50}}{50!} = \boxed{0.0273}. \quad (\text{Answer: C})$$

□

Example 1.1.3. (CAS Exam 3L Spring 2010 Question 12: Probability – II) Downloads of a song on a musician’s Web site follow a heterogeneous Poisson process with the following Poisson rate function:

$$\lambda(t) = e^{-0.25t}$$

Calculate the probability that there will be more than two downloads of this song between times $t = 1$ and $t = 5$.

- A. Less than 29%
- B. At least 29%, but less than 30%
- C. At least 30%, but less than 31%
- D. At least 31%, but less than 32%
- E. At least 32%

Solution. Because $N(5) - N(1)$ is a Poisson random variable with parameter

$$\int_1^5 \lambda(t) dt = \int_1^5 e^{-0.25t} dt = \frac{e^{-0.25(1)} - e^{-0.25(5)}}{0.25} = 1.969184,$$

the required probability equals

$$\begin{aligned} & \Pr(N(5) - N(1) > 2) \\ &= 1 - \Pr(N(5) - N(1) = 0) - \Pr(N(5) - N(1) = 1) - \Pr(N(5) - N(1) = 2) \\ &= 1 - e^{-1.969184} \left(1 + 1.969184 + \frac{1.969184^2}{2} \right) \\ &= \boxed{0.3150}. \quad (\text{Answer: D}) \end{aligned}$$

□

Example 1.1.4. (CAS Exam S Spring 2016 Question 3: Probability – III) You are given:

- The number of claims, $N(t)$, follows a Poisson process with intensity:

$$\begin{aligned} \lambda(t) &= \frac{1}{2}t, & 0 < t < 5 \\ \lambda(t) &= \frac{1}{4}t, & t \geq 5 \end{aligned}$$

- By time $t = 4$, 15 claims have occurred.

Calculate the probability that exactly 16 claims will have occurred by time $t = 6$.

- A. Less than 0.075

- B. At least 0.075, but less than 0.125
- C. At least 0.125, but less than 0.175
- D. At least 0.175, but less than 0.225
- E. At least 0.225

Solution. The number of claims between $t = 4$ and $t = 6$ is a Poisson random variable with mean

$$\int_4^5 \frac{1}{2}t dt + \int_5^6 \frac{1}{4}t dt = \frac{5^2 - 4^2}{2(2)} + \frac{6^2 - 5^2}{4(2)} = 3.625.$$

The probability of having exactly one (= $16 - 15$) claim between $t = 4$ and $t = 6$ is $3.625e^{-3.625} = \boxed{0.0966}$. (**Answer: B**) □

Remark. More formally, the probability we seek is

$$\begin{aligned} \Pr(N(6) = 16 | N(4) = 15) &= \Pr(N(6) - N(4) = 1 | N(4) = 15) \\ &= \Pr(N(6) - N(4) = 1) \end{aligned}$$

due to the property of independent increments.

Example 1.1.5. (CAS Exam 3L Spring 2012 Question 9: Expected value from now until forever) Claims reported for a group of policies follow a non-homogeneous Poisson process with rate function:

$$\lambda(t) = 100/(1+t)^3, \text{ where } t \text{ is the time (in years) after January 1, 2011.}$$

Calculate the expected number of claims reported after January 1, 2011 for this group of policies.

- A. Less than 45
- B. At least 45, but less than 55
- C. At least 55, but less than 65
- D. At least 65, but less than 75
- E. At least 75

Solution. We are interested in $N(\infty) = \lim_{t \rightarrow \infty} N(t)$, which is a Poisson random variable with mean

$$\int_0^{\infty} \lambda(t) dt = \int_0^{\infty} \frac{100}{(1+t)^3} dt = 100 \left[-\frac{1}{2(1+t)^2} \right]_0^{\infty} = \boxed{50}. \quad (\mathbf{Answer: B})$$

□

Example 1.1.6. (CAS Exam 3L Spring 2013 Question 9: Variance) You are given the following:

- An actuary takes a vacation where he will not have access to email for eight days.
- While he is away, emails arrive in the actuary's inbox following a non-homogeneous Poisson process where

$$\lambda(t) = 8t - t^2 \text{ for } 0 \leq t \leq 8. (t \text{ is in days})$$

Calculate the variance of the number of emails received by the actuary during this trip.

- A. Less than 60
- B. At least 60, but less than 70
- C. At least 70, but less than 80
- D. At least 80, but less than 90
- E. At least 90

Solution. The trip of the actuary lasts for 8 days, during which the number of emails is a Poisson random variable with variance (same as the mean)

$$\int_0^8 (8t - t^2) dt = \left[4t^2 - \frac{t^3}{3} \right]_0^8 = \boxed{85.3333}. \quad (\text{Answer: D})$$

□

Example 1.1.7. (CAS Exam ST Fall 2015 Question 1: Calculation of homogeneous Poisson intensity) For two Poisson processes, N_1 and N_2 , you are given:

- N_1 has intensity function $\lambda_1(t) = \begin{cases} 2t & \text{for } 0 < t \leq 1 \\ t^3 & \text{for } t > 1 \end{cases}$
- N_2 is a homogeneous Poisson process.
- $\text{Var}[N_1(3)] = 4\text{Var}[N_2(3)]$

Calculate the intensity of N_2 at $t = 3$.

- A. Less than 1
- B. At least 1, but less than 3
- C. At least 3, but less than 5
- D. At least 5, but less than 7
- E. At least 7

Solution. Note that $N_1(3)$ has a mean and variance equal to

$$\int_0^3 \lambda_1(t) dt = \int_0^1 2t dt + \int_1^3 t^3 dt = [t^2]_0^1 + \left[\frac{t^4}{4}\right]_1^3 = 1 + \frac{3^4 - 1^4}{4} = 21,$$

while $N_2(3)$ has a mean and variance equal to $3\lambda_2$, where λ_2 is the constant intensity of N_2 . As $\text{Var}[N_1(3)] = 4\text{Var}[N_2(3)]$, we have $21 = 4(3\lambda_2)$, so $\lambda_2 = 21/12 = \boxed{1.75}$. **(Answer: B)** \square

Probabilities and expectations involving overlapping intervals. A harder exam question may ask that you determine probabilities and expected values for increments on overlapping intervals. The key step to calculate these quantities lies in rewriting them in terms of increments on *non-overlapping* intervals, which are independent according to the definition of a Poisson process.

Example 1.1.8. (CAS Exam MAS-I Spring 2019 Question 1: First encounter with joint events) Cars arrive according to a Poisson process at a rate of two per hour.

Calculate the probability that in a given hour, exactly one car will arrive during the first ten minutes and no other cars will arrive during that hour.

- A. Less than 0.03
- B. At least 0.03 but less than 0.06
- C. At least 0.06 but less than 0.09
- D. At least 0.09 but less than 0.12
- E. At least 0.12

Solution. Let's measure time in hours. We are to find $\Pr(N(1/6) = 1, N(1) - N(1/6) = 0)$, which, by the property of independent increments, can be factored into

$$\begin{aligned} \Pr\left(N\left(\frac{1}{6}\right) = 1\right) \Pr\left(N(1) - N\left(\frac{1}{6}\right) = 0\right) &= e^{-2/6} \left(\frac{2}{6}\right) \times e^{-2(1-1/6)} \\ &= \boxed{0.0451}. \quad \textbf{(Answer: B)} \end{aligned}$$

\square

Example 1.1.9. (Probability for overlapping increments) The number of calls received by a call center follow a homogeneous Poisson process with a rate of 30 per hour.

Calculate the probability that there are exactly 2 calls in the first ten minutes and exactly 5 calls in the first twenty minutes.

- A. Less than 0.01
- B. At least 0.01, but less than 0.02

- C. At least 0.02, but less than 0.03
- D. At least 0.03, but less than 0.04
- E. At least 0.04

Solution. When time is measured in hours, the required probability is

$$\Pr(N(1/6) = 2, N(1/3) = 5) = \Pr(N(1/6) = 2, N(1/3) - N(1/6) = 3),$$

which can be factored, because of independence, into

$$\begin{aligned} \Pr(N(1/6) = 2) \Pr(N(1/3) - N(1/6) = 3) &= \frac{e^{-30/6}(30/6)^2}{2!} \times \frac{e^{-30/6}(30/6)^3}{3!} \\ &= \boxed{0.0118}. \quad (\text{Answer: B}) \end{aligned}$$

□

Example 1.1.10. (CAS Exam MAS-I Fall 2018 Question 1: Expectation of a product)

Insurance claims are made according to a Poisson process $\{N(t), t \geq 0\}$ with rate $\lambda = 1$.

Calculate $E[N(1) * N(2)]$.

- A. Less than 1.5
- B. At least 1.5, but less than 2.5
- C. At least 2.5, but less than 3.5
- D. At least 3.5, but less than 4.5
- E. At least 4.5

Solution. Writing $N(2)$ as the telescoping sum $N(2) = (N(2) - N(1)) + N(1)$, we get

$$\begin{aligned} E[N(1)N(2)] &= E\{N(1) \times [(N(2) - N(1)) + N(1)]\} \\ &= E[N(1)(N(2) - N(1))] + E[N(1)^2]. \end{aligned}$$

By the independent increments property, $N(2) - N(1)$ is independent of $N(1)$, so

$$E[N(1)(N(2) - N(1))] = E[N(1)]E[N(2) - N(1)] = 1 \times (2 - 1) = 1.$$

Moreover,

$$E[N(1)^2] = \text{Var}[N(1)] + E[N(1)]^2 = 1 + 1^2 = 2.$$

Finally,

$$E[N(1)N(2)] = 1 + 2 = \boxed{3}. \quad (\text{Answer: C})$$

□

Remark. Of course, $E[N(1)N(2)] \neq E[N(1)]E[N(2)] = 2$, which corresponds to Answer B.

Conditional distribution of $N(t)$ given $N(s)$ for $s \leq t$. Suppose that we know the value of the Poisson process at one time point s with $N(s) = m$, and we wish to study the probabilistic behavior of the Poisson process at a later time point t with $s \leq t$. Then $N(t)$ turns out to be a *translated Poisson* random variable in the sense that it has the same distribution as the sum of a Poisson random variable and a constant. To see this, let's write $N(t)$ as

$$N(t) = [N(t) - N(s)] + N(s).$$

The second term $N(s)$ is known to be m , while the first term, owing to the property of independent increments of a Poisson process, is a Poisson random variable, say M , whose distribution does not depend on the value of m . Therefore, we have the distributional representation

$$\boxed{[N(t)|N(s) = m] \stackrel{d}{=} M + m, \quad s \leq t,}$$

where “ $\stackrel{d}{=}$ ” means equality in distribution. This result allows us to answer questions about many probabilistic quantities associated with $N(t)$ when the value of $N(s)$ is given.

Example 1.1.11. (CAS Exam S Fall 2017 Question 3: Conditional probability) You are given:

- A Poisson process N has a rate function: $\lambda(t) = 3t^2$
- You've already observed 50 events by time $t = 2.1$.

Calculate the conditional probability, $\Pr[N(3) = 68 | N(2.1) = 50]$.

- A. Less than 5%
- B. At least 5%, but less than 10%
- C. At least 10%, but less than 15%
- D. At least 15%, but less than 20%
- E. At least 20%

Solution. The conditional probability can be determined as

$$\begin{aligned} \Pr[N(3) = 68 | N(2.1) = 50] &= \Pr[N(3) - N(2.1) = 18 | N(2.1) = 50] \\ &= \Pr[N(3) - N(2.1) = 18], \end{aligned}$$

where $N(3) - N(2.1)$ is a Poisson random variable with mean $\int_{2.1}^3 3t^2 dt = 3^3 - 2.1^3 = 17.739$. The final answer is

$$\frac{e^{-17.739} 17.739^{18}}{18!} = \boxed{0.0934}. \quad (\text{Answer: B})$$

□

Example 1.1.12. (CAS Exam 3L Fall 2010 Question 11: Conditional variance) You are given the following information:

- A Poisson process N has a rate function $\lambda(t) = 3t^2$.
- You have observed 50 events by time $t = 2.1$.

Calculate $\text{Var}[N(3) \mid N(2.1) = 50]$.

- A. Less than 10
- B. At least 10, but less than 20
- C. At least 20, but less than 30
- D. At least 30, but less than 40
- E. At least 40

Solution. Conditional on $N(2.1) = 50$, $N(3)$ has the same distribution as $M + 50$, where M is a Poisson random variable with mean and variance

$$\int_{2.1}^3 \lambda(t) dt = \int_{2.1}^3 3t^2 dt = t^3 \Big|_{2.1}^3 = 17.739.$$

Hence

$$\text{Var}[N(3) \mid N(2.1) = 50] = \text{Var}(M + 50) = \text{Var}(M) = \boxed{17.739} \quad \text{(Answer: B)}.$$

□

Normal approximation. For more cumbersome probabilities such as $\Pr(N(t) > c)$ with c being a large number, exact calculations can be tedious and normal approximation may be used. That is, we approximate $N(t)$ by a normal random variable with the same mean and variance, and instead calculate the probability of the same event for this normal random variable using the standard normal distribution table you have in the exam. Because the distribution of $N(t)$ is discrete, and a continuous distribution (i.e., normal) is used to approximate this discrete distribution, a continuity correction should be made.

Recall - Continuity correction

Let X be a random variable taking values in the set of integers $\{0, \pm 1, \pm 2, \dots\}$ and N is a normal random variable having the same mean and variance as X . The following shows how various probabilities are approximated using the normal approximation *with* continuity correction: (c is an integer)

Probability of Interest		Approximant
$\Pr(X \leq c)$	\approx	$\Pr(N \leq c + 0.5)$
$\Pr(X < c)$	\approx	$\Pr(N \leq c - 0.5)$
$\Pr(X \geq c)$	\approx	$\Pr(N \geq c - 0.5)$
$\Pr(X > c)$	\approx	$\Pr(N > c + 0.5)$

In the second column, it does not matter whether we take strict or weak inequalities because N is a continuous random variable. In other words, we may replace " \leq " by " $<$ " and " \geq " by " $>$ ".

Throughout this study manual, we denote the distribution function of the standard normal distribution by $\Phi(\cdot)$. (You may have used $N(\cdot)$ in Exam MFE/IFM/3F)

Example 1.1.13. (CAS Exam 3L Spring 2008 Question 11: Normal approximation)

A customer service call center operates from 9:00 AM to 5:00 PM. The number of calls received by the call center follows a Poisson process whose rate function varies according to the time of day, as follows:

Time of Day	Call Rate (per hour)
9:00 AM to 12:00 PM	30
12:00 PM to 1 :00 PM	10
1:00 PM to 3:00 PM	25
3:00 PM to 5:00 PM	30

Using a normal approximation, what is the probability that the number of calls received from 9:00AM to 1:00PM exceeds the number of calls received from 1:00PM to 5:00PM?

- A. Less than 10%
- B. At least 10%, but less than 20%
- C. At least 20%, but less than 30%
- D. At least 30%, but less than 40%
- E. At least 40%

Solution. The number of calls received from 9:00AM to 1:00PM is a Poisson random variable N^1 with parameter $30(3) + 10(1) = 100$, while the number of calls received from 1:00PM to

5:00PM is a Poisson random variable N^2 with parameter $25(2) + 30(2) = 110$. Because N^1 and N^2 are independent,

$$E[N^1 - N^2] = E[N^1] - E[N^2] = 100 - 110 = -10,$$

and

$$\text{Var}(N^1 - N^2) = \text{Var}(N^1) + \text{Var}(N^2) = 100 + 110 = 210.$$

Using the normal approximation with continuity correction, we have

$$\Pr(N^1 > N^2) = \Pr(N^1 - N^2 > 0) \approx \Pr\left(\underbrace{N(-10, 210)}_{\substack{\text{a normal r.v. with mean} \\ -10 \text{ and variance } 210}} > 0.5\right),$$

which, upon standardization, equals

$$\Pr\left(N(0, 1) > \frac{0.5 - (-10)}{\sqrt{210}}\right) = 1 - \Phi(0.72) = 1 - 0.7642 = \boxed{0.2358}. \quad (\text{Answer: C})$$

□

Remark. If you do not use continuity correction, you will calculate

$$\Pr(N^1 > N^2) \approx \Pr(N(-10, 210) > 0) = 1 - \Phi\left(\frac{0 - (-10)}{\sqrt{210}}\right) = 1 - \underbrace{\Phi(0.69)}_{0.7549} = 0.2451,$$

in which case you will also end up with Answer C.

[HARDER!] Conditional distribution of $N(s)$ given $N(t)$ with $s \leq t$. We have learned that conditional on $N(s)$, the distribution of $N(t)$, where $0 \leq s \leq t$, is that of a translated Poisson distribution. What about the conditional distribution of $N(s)$ given $N(t)$? To answer this question, we consider, for $k = 0, 1, \dots, n$,

$$\Pr(N(s) = k | N(t) = n) = \frac{\Pr(N(s) = k, N(t) = n)}{\Pr(N(t) = n)} = \frac{\Pr(N(s) = k, N(t) - N(s) = n - k)}{\Pr(N(t) = n)}.$$

Because a (homogeneous or non-homogeneous) Poisson process possesses independent increments, the preceding probability can be further written as

$$\begin{aligned} \Pr(N(s) = k | N(t) = n) &= \frac{\Pr(N(s) = k) \Pr(N(t) - N(s) = n - k)}{\Pr(N(t) = n)} \\ &= \frac{e^{-m(s)} [m(s)]^k / k! \times e^{-[m(t) - m(s)]} [m(t) - m(s)]^{n-k} / (n-k)!}{e^{-m(t)} [m(t)]^n / n!} \\ &= \frac{n!}{k!(n-k)!} \left(\frac{m(s)}{m(t)}\right)^k \left(1 - \frac{m(s)}{m(t)}\right)^{n-k} \\ &= \binom{n}{k} \left(\frac{m(s)}{m(t)}\right)^k \left(1 - \frac{m(s)}{m(t)}\right)^{n-k}. \end{aligned}$$

In other words, given $N(t) = n$, $N(s)$ is a binomial random variable with parameters n and $m(s)/m(t)$. In particular, for a homogeneous Poisson process with rate λ , i.e., $m(t) = \lambda t$ for $t \geq 0$, then

$$N(s)|N(t) = n \sim \text{Binomial}\left(n, \frac{s}{t}\right),$$

which is free of λ .

Example 1.1.14. (CAS Exam MAS-I Spring 2018 Question 2: Probability for $N(s)$ given $N(t)$ with $s \leq t$) Insurance claims are made according to a Poisson process with rate λ .

Calculate the probability that exactly 3 claims were made by time $t = 1$, given that exactly 6 claims are made by time $t = 2$.

- A. Less than 0.3
- B. At least 0.3, but less than 0.4
- C. At least 0.4, but less than 0.5
- D. At least 0.5, but less than 0.6
- E. At least 0.6

Solution. Conditional on $N(2) = 6$, $N(1)$ is a binomial random variable with parameters 6 and $1/2$, so

$$\Pr[N(1) = 3|N(2) = 6] = \binom{6}{3} \left(\frac{1}{2}\right)^3 \left(1 - \frac{1}{2}\right)^3 = \boxed{0.3125}. \quad (\text{Answer: B})$$

□

Remark. We do not need the value of λ to get the answer.

Problems

Note: In general, end-of-chapter/section problems in this study manual are listed in the same order as they are discussed in the main text of each chapter/section. Don't be scared by the abundance of problems. As a general guide it would be a good idea to work out *at least all of the past CAS problems*. Hard work pays!

Problem 1.1.1. (SOA Exam P Sample Question 280: Conditional Poisson mean) The number of burglaries occurring on Burlington Street during a one-year period is Poisson distributed with mean 1.

Calculate the expected number of burglaries on Burlington Street in a one-year period, given that there are at least two burglaries.

- A. 0.63
- B. 2.39
- C. 2.54
- D. 3.00
- E. 3.78

Solution. Let N be the number of burglaries on Burlington Street in a specified one-year period. Given that there are at least two burglaries, the expected value of N is

$$\begin{aligned}
 E[N|N \geq 2] &= \frac{\sum_{n=2}^{\infty} n \Pr(N = n)}{\Pr(N \geq 2)} \\
 &= \frac{E[N] - 1 \times \Pr(N = 1)}{1 - \Pr(N = 0) - \Pr(N = 1)} \\
 &= \frac{1 - e^{-1}}{1 - e^{-1} - e^{-1}} \\
 &= \boxed{2.3922}. \quad (\text{Answer: B})
 \end{aligned}$$

□

Problem 1.1.2. (CAS Exam 3 Fall 2006 Question 26: True-of-false questions) Which of the following is/are true?

1. A counting process is said to possess independent increments if the number of events that occur between time s and t is independent of the number of events that occur between time s and $t + u$ for all $u > 0$.
2. All Poisson processes have stationary and independent increments.
3. The assumption of stationary and independent increments is essentially equivalent to asserting that at any point in time the process probabilistically restarts itself.

- A. 1 only
- B. 2 only

- C. 3 only
- D. 1 and 2 only
- E. 2 and 3 only

Solution. Only 3. is correct. **(Answer: C)**

1. This would be true if the second s is changed to t .
2. A non-homogeneous Poisson process does not have stationary increments in general.

□

Problem 1.1.3. (CAS Exam 3L Fall 2013 Question 9: Polynomial intensity function)

You are given that claim counts follow a non-homogeneous Poisson Process with $\lambda(t) = 30t^2 + t^3$.

Calculate the probability of at least two claims between time 0.2 and 0.3.

- A. Less than 1%
- B. At least 1%, but less than 2%
- C. At least 2%, but less than 3%
- D. At least 3%, but less than 4%
- E. At least 4%

Solution. The number of claims between times 0.2 and 0.3 is a Poisson random variable with parameter

$$\int_{0.2}^{0.3} (30t^2 + t^3) dt = 10t^3 + \frac{t^4}{4} \Big|_{0.2}^{0.3} = 0.191625.$$

Hence the probability of at least two claims between times 0.2 and 0.3 is the complement of the probability of having 0 or 1 claim:

$$1 - \Pr(0 \text{ claim}) - \Pr(1 \text{ claim}) = 1 - e^{-0.191625}(1 + 0.191625) = \boxed{0.0162}. \quad \text{(Answer: B)}$$

□

Problem 1.1.4. (CAS Exam 3 Fall 2006 Question 28: Piecewise linear intensity function)

Customers arrive to buy lemonade according to a Poisson distribution with $\lambda(t)$, where t is time in hours, as follows:

$$\lambda(t) = \begin{cases} 2 + 6t & 0 \leq t \leq 3 \\ 20 & 3 < t \leq 4 \\ 36 - 4t & 4 < t \leq 8 \end{cases}$$

At 9:00 a.m., t is 0.

Calculate the number of customers expected to arrive between 10:00 a.m. and 2:00 p.m.

- A. Less than 63
- B. At least 63, but less than 65

- C. At least 65, but less than 67
- D. At least 67, but less than 69
- E. At least 69

Solution. The expected number of customers arriving between 10:00 a.m. ($t = 1$) and 2:00 p.m. ($t = 5$) is

$$\begin{aligned} \int_1^5 \lambda(t) dt &= \int_1^3 (2 + 6t) dt + \int_3^4 20 dt + \int_4^5 (36 - 4t) dt \\ &= [2t + 3t^2]_1^3 + 20 + [36t - 2t^2]_4^5 \\ &= \boxed{66}. \quad (\text{Answer: C}) \end{aligned}$$

□

Remark. Because the intensity function is piecewise linear, integrating it is the same as calculating the areas of trapeziums.

Problem 1.1.5. (SOA Course 3 Fall 2004 Question 26: Linear rate function) Customers arrive at a store at a Poisson rate that increases linearly from 6 per hour at 1:00 p.m. to 9 per hour at 2:00 p.m.

Calculate the probability that exactly 2 customers arrive between 1:00 p.m. and 2:00 p.m.

- A. 0.016
- B. 0.018
- C. 0.020
- D. 0.022
- E. 0.024

Solution. Let 1:00 p.m. be time 0 and measure time in hours. The rate function is given by

$$\lambda(t) = 6 + 3t, \quad t \geq 0.$$

You can check that $\lambda(0) = 6$ and $\lambda(1) = 9$. The number of customers that arrive between 1:00 p.m. and 2:00 p.m. is a Poisson random variable with a mean of $\int_0^1 \lambda(t) dt = 6 + 3/2 = 7.5$. The probability of having 2 customers in the same period is

$$\frac{e^{-7.5}(7.5)^2}{2!} = \boxed{0.0156}. \quad (\text{Answer: A})$$

□

Problem 1.1.6. (CAS Exam 3L Fall 2008 Question 1: Expected value) The number of accidents on a highway from 3:00 PM to 7:00 PM follows a nonhomogeneous Poisson process with rate function

$$\lambda = 4 - (t - 2)^2, \quad \text{where } t \text{ is the number of hours since 3:00 PM.}$$

How many more accidents are expected from 4:00 PM to 5:00 PM than from 3:00 PM to 4:00PM?

- A. Less than 0.75
- B. At least 0.75, but less than 1.25
- C. At least 1.25, but less than 1.75
- D. At least 1.75, but less than 2.25
- E. At least 2.25

Solution. • The expected number of accidents from 3:00 PM to 4:00PM is

$$\int_0^1 [4 - (t - 2)^2] dt = \left[4t - \frac{(t - 2)^3}{3} \right]_0^1 = \frac{5}{3}.$$

- The expected number of accidents from 4:00 PM to 5:00 PM is

$$\int_1^2 [4 - (t - 2)^2] dt = \left[4t - \frac{(t - 2)^3}{3} \right]_1^2 = \frac{11}{3}.$$

The difference is $\boxed{2}$. (**Answer: D**) □

Problem 1.1.7. (CAS Exam 3L Fall 2008 Question 2: Probability, homogeneous) You are given the following:

- Hurricanes occur at a Poisson rate of $1/4$ per week during the hurricane season.
- The hurricane season lasts for exactly 15 weeks.

Prior to the next hurricane season, a weather forecaster makes the statement, “There will be at least three and no more than five hurricanes in the upcoming hurricane season.”

Calculate the probability that this statement will be correct.

- A. Less than 54%
- B. At least 54%, but less than 56%
- C. At least 56%, but less than 58%
- D. At least 58%, but less than 60%
- E. At least 60%

Solution. Note that $N(15)$, the number of hurricanes during the 15-week hurricane season, is a Poisson random variable with a mean of $15/4 = 3.75$. The probability that the statement will be correct is

$$\Pr(3 \leq N(15) \leq 5) = e^{-3.75} \left(\frac{3.75^3}{3!} + \frac{3.75^4}{4!} + \frac{3.75^5}{5!} \right) = \boxed{0.5458}. \quad (\mathbf{Answer: B})$$

□

Problem 1.1.8. (CAS Exam 3L Spring 2008 Question 10: Probability, non-homogeneous) Car accidents follow a Poisson process, as described below:

- On Monday and Friday, the expected number of accidents per day is 3.
- On Tuesday, Wednesday, and Thursday, the expected number of accidents per day is 4.
- On Saturday and Sunday, the expected number of accidents per day is 1.

Calculate the probability that exactly 18 accidents occur in a week.

- A. Less than .06
- B. At least .06 but less than .07
- C. At least .07 but less than .08
- D. At least .08 but less than .09
- E. At least .09

Solution. The total number of accidents in a week is a Poisson random variable with a mean of $3(2) + 4(3) + 1(2) = 20$, so the probability of having exactly 18 accidents in a week is

$$\frac{e^{-20}20^{18}}{18!} = \boxed{0.0844}. \quad (\text{Answer: D})$$

□

Problem 1.1.9. (CAS Exam 3 Spring 2006 Question 33: Probability, non-homogeneous)

While on vacation, an actuarial student sets out to photograph a Jackalope and a Snipe, two animals common to the local area. A tourist information booth informs the student that daily sightings of Jackalopes and Snipes follow independent Poisson processes with intensity parameters:

$$\begin{aligned} \lambda_J(t) &= \frac{t^{1/3}}{5} && \text{for Jackalopes} \\ \lambda_S(t) &= \frac{t^{1/2}}{10} && \text{for Snipes} \end{aligned}$$

where: $0 \leq t \leq 24$ and t is the number of hours past midnight

If the student takes photographs between 1 pm and 5 pm, calculate the probability that he will take at least 1 photograph of each animal.

- A. Less than 0.45
- B. At least 0.45, but less than 0.60
- C. At least 0.60, but less than 0.75
- D. At least 0.75, but less than 0.90
- E. At least 0.90

Solution. The number of Jackalopes and Snipes between 1 pm and 5 pm are Poisson random variables with respective means

$$\frac{1}{5} \int_{13}^{17} t^{1/3} dt = \frac{3}{4(5)} (17^{4/3} - 13^{4/3}) = 1.971665$$

and

$$\frac{1}{10} \int_{13}^{17} t^{1/2} dt = \frac{2}{3(10)} (17^{3/2} - 13^{3/2}) = 1.548042.$$

Because the two Poisson processes are independent (note: here we are not using the property of independent increments),

$$\begin{aligned} & \Pr(N_J(17) - N_J(13) \geq 1, N_S(17) - N_S(13) \geq 1) \\ &= \Pr(N_J(17) - N_J(13) \geq 1) \Pr(N_S(17) - N_S(13) \geq 1) \\ &= [1 - \Pr(N_J(17) - N_J(13) = 0)] [1 - \Pr(N_S(17) - N_S(13) = 0)] \\ &= (1 - e^{-1.971665})(1 - e^{-1.548042}) \\ &= \boxed{0.6777}. \quad \text{(Answer: C)} \end{aligned}$$

□

Problem 1.1.10. (CAS Exam 3 Fall 2005 Question 26: Probability, non-homogeneous)

The number of reindeer injuries on December 24 follows a Poisson process with intensity function:

$$\lambda(t) = (t/12)^{1/2} \quad 0 \leq t \leq 24, \text{ where } t \text{ is measured in hours}$$

Calculate the probability that no reindeer will be injured during the last hour of the day.

- A. Less than 30%
- B. At least 30%, but less than 40%
- C. At least 40%, but less than 50%
- D. At least 50%, but less than 60%
- E. At least 60%

Solution. We need

$$\begin{aligned} \Pr(N(24) - N(23) = 0) &= \exp \left[- \int_{23}^{24} (t/12)^{1/2} dt \right] \\ &= \exp \left[- \frac{2}{3(12)^{1/2}} (24^{3/2} - 23^{3/2}) \right] \\ &= \boxed{0.24675}. \quad \text{(Answer: A)} \end{aligned}$$

□

Problem 1.1.11. (Probability for overlapping increments) Customers arrive at a post office in accordance with a Poisson process with a rate of 5 per hour. The post office opens at 9:00 am.

Calculate the probability that only one customer arrives before 9:20 am and ten customers arrive before 11:20 am.

- A. Less than 0.01
- B. At least 0.01, but less than 0.02
- C. At least 0.02, but less than 0.03
- D. At least 0.03, but less than 0.04
- E. At least 0.04

Solution. The required probability is

$$\begin{aligned}
 \Pr(N(1/3) = 1, N(7/3) = 10) &= \Pr(N(1/3) = 1, N(7/3) - N(1/3) = 9) \\
 &= \Pr(N(1/3) = 1) \Pr(N(7/3) - N(1/3) = 9) \\
 &= e^{-5/3} \binom{5}{3} \times e^{-5(2)} \frac{[5(2)]^9}{9!} \\
 &= \boxed{0.0394}. \quad (\text{Answer: D})
 \end{aligned}$$

□

Problem 1.1.12. [HARDER!] (Rate function mimics the normal density function) You are given that claim counts follow a non-homogeneous Poisson process with intensity function $\lambda(t) = e^{-t^2/4}$.

Calculate the probability of at least two claims between time 1 and time 2.

- A. Less than 0.10
- B. At least 0.10, but less than 0.15
- C. At least 0.15, but less than 0.20
- D. At least 0.20, but less than 0.25
- E. At least 0.25

Solution. The number of claims between time 1 and time 2 is a Poisson random variable with mean

$$\int_1^2 \lambda(t) dt = \int_1^2 e^{-t^2/4} dt.$$

Note that the integrand resembles the density function of a normal distribution with mean 0 and variance 2, except that the normalizing constant $1/\sqrt{2\pi(2)}$ is missing. Hence

$$\begin{aligned}
 \int_1^2 e^{-t^2/4} dt &= \sqrt{2\pi(2)} \int_1^2 \frac{1}{\sqrt{2\pi(2)}} e^{-t^2/4} dt \\
 &= 2\sqrt{\pi} \Pr(1 < N(0, 2) < 2) \\
 &= 2\sqrt{\pi} \left[\Phi\left(\frac{2}{\sqrt{2}}\right) - \Phi\left(\frac{1}{\sqrt{2}}\right) \right] \\
 &= 2\sqrt{\pi} [\Phi(1.41) - \Phi(0.71)] \\
 &= 2\sqrt{\pi} (0.9207 - 0.7611) \\
 &= 0.565767.
 \end{aligned}$$

Finally, the required probability is

$$\begin{aligned}\Pr(N(2) - N(1) \geq 2) &= 1 - \Pr(N(2) - N(1) \leq 1) \\ &= 1 - e^{-0.565767}(1 + 0.565767) \\ &= \boxed{0.1108}. \quad (\text{Answer: B})\end{aligned}$$

□

Problem 1.1.13. (CAS Exam ST Spring 2016 Question 1: Probability for $N(s)$ given $N(t)$ with $s \leq t$ – I) You are given that $N(t)$ follows the Poisson process with rate $\lambda = 2$.

Calculate $\Pr[N(2) = 3 | N(5) = 7]$.

- A. Less than 0.25
- B. At least 0.25, but less than 0.35
- C. At least 0.35, but less than 0.45
- D. At least 0.45, but less than 0.55
- E. At least 0.55

Solution. Conditional on $N(5) = 7$, $N(2)$ is a binomial random variable with parameters 7 and $2/5$, so

$$\Pr[N(2) = 3 | N(5) = 7] = \binom{7}{3} \left(\frac{2}{5}\right)^3 \left(1 - \frac{2}{5}\right)^4 = \boxed{0.290304}. \quad (\text{Answer: B})$$

□

Remark. We do not need the value of λ to get the answer.

Problem 1.1.14. (Probability for $N(s)$ given $N(t)$ for $s \leq t$ – II) Customers arrive at a post office in accordance with a Poisson process with a rate of 5 per hour. The post office opens at 9:00 am.

Ten customers have arrived before 11:00 am.

Calculate the probability that only two customers have arrived before 9:30 am.

- A. Less than 0.15
- B. At least 0.15, but less than 0.20
- C. At least 0.20, but less than 0.25
- D. At least 0.25, but less than 0.30
- E. At least 0.30

Solution. Note that $N(0.5) | N(2) = 10$ has a binomial distribution with parameters 10 and $0.5/2 = 0.25$. The conditional probability that $N(0.5) = 2$ equals

$$\binom{10}{2} 0.25^2 (1 - 0.25)^8 = \boxed{0.2816}. \quad (\text{Answer: D})$$

□

Problem 1.1.15. [HARDER!] (Mean of a conditional sandwiched Poisson process value)

You are given that $\{N(t)\}$ is a Poisson process with rate $\lambda = 2$.

Calculate the expected value of $N(3)$, conditional on $N(2) = 3$ and $N(5) = 10$.

- A. Less than 5
- B. At least 5, but less than 6
- C. At least 6, but less than 7
- D. At least 7, but less than 8
- E. At least 8

Solution. We are interested in the distribution of the value of a Poisson process at a particular time point, given the process values at an *earlier* time as well as a *later* time. To reduce this two-condition setting to the one-condition setting involving $N(s)$ given $N(t)$ for $s \leq t$, we consider the *translated Poisson process* $\{N^2(t)\}_{t \geq 0}$ defined by $N^2(t) := N(2+t) - N(2)$. This translated process looks at the original Poisson process $\{N(t)\}$ from time 2 (hence the superscript “2”) onward, but with values translated downward by $N(2)$ units. It is easy to conceive (and can be rigorously shown) that $\{N^2(t)\}_{t \geq 0}$ is indeed a Poisson process (see Exercise 5.35 of Ross). Moreover, because of independent increments, $\{N^2(t)\}_{t \geq 0}$ is independent of $N(2)$. In terms of the translated Poisson process, $N(3)$ can be written as the telescoping sum

$$N(3) = [N(3) - N(2)] + N(2) = N^2(1) + 3.$$

As $\{N(2) = 3, N(5) = 10\} = \{N(2) = 3, N^2(3) = 7\}$, the conditional distribution of $N^2(1)$ is

$$N^2(1) \mid \underbrace{[N(2) = 3, N^2(3) = 7]}_{\text{get rid of this}} \sim N^2(1) \mid N^2(3) = 7 \sim \text{Bin}(7, 1/3).$$

Finally, the conditional expected value of $N(3)$ is

$$\begin{aligned} \mathbb{E}[N(3) \mid N(2) = 3, N(5) = 10] &= 3 + \mathbb{E}[N^2(1) \mid N(2) = 3, N^2(3) = 7] \\ &= 3 + \frac{7}{3} \\ &= \boxed{5.3333}. \quad (\text{Answer: B}) \end{aligned}$$

□

Remark. In general, for $t_1 \leq s \leq t_2$ and a non-homogeneous Poisson process with mean value function $m(\cdot)$, the distribution of $N(s)$ conditional on $N(t_1) = A$ and $N(t_2) = B$ is

$$A + \text{Bin}\left(B - A, \frac{m(s) - m(t_1)}{m(t_2) - m(t_1)}\right).$$

The conditional expected value is

$$A + (B - A) \left[\frac{m(s) - m(t_1)}{m(t_2) - m(t_1)} \right] = \left[\frac{m(t_2) - m(s)}{m(t_2) - m(t_1)} \right] A + \left[\frac{m(s) - m(t_1)}{m(t_2) - m(t_1)} \right] B,$$

which is a weighted average of A and B .